

日本国特許庁
JAPAN PATENT OFFICE

#2 8-8-01
Priority Paper

JC821 U.S. PRO
09/898639



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日
Date of Application:

2000年 7月 5日

出願番号
Application Number:

特願2000-203157

出願人
Applicant(s):

日本電気株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 5月11日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



出証番号 出証特2001-3040385

【書類名】 特許願

【整理番号】 68501845

【提出日】 平成12年 7月 5日

【あて先】 特許庁長官殿

【国際特許分類】 H03M 7/30

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 長谷川 聡

【発明者】

 【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

 【氏名】 高見沢 雄一郎

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100088812

 【弁理士】

 【氏名又は名称】 ▲柳▼川 信

【手数料の表示】

 【予納台帳番号】 030982

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声符号化装置及びそれに用いる心理聴覚分析方法

【特許請求の範囲】

【請求項1】 入力信号を複数の周波数帯域に分割する分割手段を持ち、前記分割手段で分割された各サブバンド信号を圧縮符号化する音声符号化装置であって、前記各サブバンド信号の各周波数について聴感上の音の大きさが等しい音圧レベルの値を結んだ等ラウドネス曲線に準拠した重み付けを行いかつその重み付けされた量子化誤差が前記各サブバンド信号で均等になるようにビット割当てを行う手段を有することを特徴とする音声符号化装置。

【請求項2】 前記各サブバンド信号に対して前記等ラウドネス曲線に準拠した重み付け係数を保持するテーブルを前記ビット割当てを行う手段に含むことを特徴とする請求項1記載の音声符号化装置。

【請求項3】 符号化ビットレートに対応した重み付け係数を保持する重み付けテーブルを前記ビット割当てを行う手段に含み、前記符号化ビットレートに応じた重み付けが付加された量子化誤差が前記各サブバンド信号で均等になるように前記ビット割当てを行うようにしたことを特徴とする請求項2記載の音声符号化装置。

【請求項4】 前記重み付けテーブルを前記符号化ビットレートに応じて複数個設け、その複数の重み付けテーブルのうちの使用するテーブルを適宜換えるようにしたことを特徴とする請求項3記載の音声符号化装置。

【請求項5】 音声符号化方式は、人間の聴覚能力の限界やマスキング効果等の聴覚特性を考慮した心理聴覚分析を用いる符号化方式であることを特徴とする請求項1から請求項4のいずれか記載の音声符号化装置。

【請求項6】 入力信号を複数の周波数帯域に分割する分割手段を持ち、前記分割手段で分割された各サブバンド信号を圧縮符号化する音声符号化装置の心理聴覚分析方法であって、前記各サブバンド信号の各周波数について聴感上の音の大きさが等しい音圧レベルの値を結んだ等ラウドネス曲線に準拠した重み付けを行いかつその重み付けされた量子化誤差が前記各サブバンド信号で均等になる

ようにビット割当てを行うステップを有することを特徴とする心理聴覚分析方法

【請求項 7】 前記ビット割当てを行うステップは、前記各サブバンド信号に対して前記等ラウドネス曲線に準拠した重み付け係数を保持するテーブルの保持内容を基に前記ビット割当てを行うようにしたことを特徴とする請求項 6 記載の心理聴覚分析方法。

【請求項 8】 前記ビット割当てを行うステップは、符号化ビットレートに対応した重み付け係数を保持する重み付けテーブルの保持内容を基に、前記符号化ビットレートに応じた重み付けが付加された量子化誤差が前記各サブバンド信号で均等になるように前記ビット割当てを行うようにしたことを特徴とする請求項 7 記載の心理聴覚分析方法。

【請求項 9】 前記重み付けテーブルを前記符号化ビットレートに応じて複数個設け、その複数の重み付けテーブルのうちの使用するテーブルを適宜換えるようにしたことを特徴とする請求項 8 記載の心理聴覚分析方法。

【請求項 10】 人間の聴覚特性を考慮する音声符号化方式に適用可能としたことを特徴とする請求項 6 から請求項 9 のいずれか記載の心理聴覚分析方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は音声符号化装置及びそれに用いる心理聴覚分析方法に関し、特に M P E G (M o v i n g P i c t u r e E x p e r t s G r o u p) 方式のように人間の聴覚心理を利用した音声符号化処理に関する。

【0002】

【従来の技術】

パーソナルコンピュータ等の情報処理装置に搭載された C P U (中央処理装置) 上で動作するソフトウェアにおいては、 M P E G 方式のように人間の聴覚心理を利用した音声符号化処理を実現する場合、一般に聴覚心理モデルと呼ばれる人間の聴覚能力の限界やマスキング効果を計算する部分の処理負荷が非常に重くなっている。

【 0 0 0 3 】

そのため、動作させる装置の性能によっては、特に実時間符号化（リアルタイム符号化）処理を施した場合に、符号化処理が間に合わずに、復号時に音声途切れが生じてしまうことがある。

【 0 0 0 4 】

上記の処理に用いられるMPEG1/Audioレイヤ1方式による音声符号化処理装置の構成を図8に示す。図8を参照すると、符号化装置2はサブバンド分析部21と、スケーリング部22と、ビット割当て部23と、量子化部24と、ビットストリーム生成部25と、聴覚心理モデルを使用した心理聴覚分析部26とから構成されている。

【 0 0 0 5 】

サブバンド分析部21は入力信号を複数の周波数帯域に分割する。スケーリング部22は各サブバンド信号に対して基準値からの倍率であるスケールファクタを計算し、ダイナミックレンジを揃える。

【 0 0 0 6 】

心理聴覚分析部26は各サブバンドで音声信号がマスキングされている比率を求める。ビット割当て部23はその心理聴覚分析部26からの結果を基に各サブバンドへのビット割当てを行う。量子化部24は量子化計算を行う。ビットストリーム生成部25はヘッダや補助情報と共にビット列を形成する。

【 0 0 0 7 】

上記の心理聴覚分析部26の構成を図10に示す。図10を参照すると、心理聴覚分析部26はFFT（高速フーリエ変換）部31と、スペクトル検出部32と、マスキングしきい値計算部33と、信号対マスク比算出部34と、音圧レベル算出部35とから構成されている。

【 0 0 0 8 】

この心理聴覚分析部26において、入力音声データをFFT部31でスペクトル分解し、このスペクトルのうち、マスカーとなり得るスペクトルのみをスペクトル検出部32で検出する。マスキングしきい値計算部33ではスペクトル検出部32で検出されたスペクトルに対し、最小可聴しきい値との比較や、マスキン

グ効果の分析を施した後、各サブバンド当たりのマスキング量を算出する。

【0009】

最終的に、音圧レベル算出部24で算出された各サブバンド当たりの音圧レベルとマスキング量とから信号対マスク比（SMR）として信号対マスク比算出部34からビット割当て部23に対して出力される。

【0010】

また、ビット割当て部23の動作フローを図9を用いて説明する。各サブバンドの量子化ステップ値を“0”に初期化し（図9ステップS31）、各サブバンドに対するマスク対ノイズ比（MNR）を算出する（図9ステップS32）。

【0011】

このうちの最小のMNRを持つサブバンドに対して量子化ステップ値を1段階増加させた後（図9ステップS33）、MNRを更新する（図9ステップS34）。ここで、現在までに割当てられている総符号量を求め（図9ステップS35）、許容符号量との比較をする。

【0012】

許容符号量に達していない場合には（図9ステップS36）、再びステップS33に戻り、ビット割当て処理を継続する。一方、許容符号量に達した場合には（図9ステップS36）、ビット割当て処理を終了する。

【0013】

【発明が解決しようとする課題】

上述した従来の音声符号化処理では、一般に聴覚心理モデルと呼ばれる人間の聴覚能力の限界やマスキング効果を計算する部分の処理負荷が重いことに加え、ビット割当て処理においてビット割当て優先順位の高いサブバンドから順にビットを割当てることから、繰り返し処理によるループ回数が多くなり、処理負荷が重くなるという問題がある。

【0014】

上記の音声符号化処理以外にも以下のような音声符号化処理方法がある。特開平10-304360号公報には音声符号化処理の負荷軽減方法が記述されており、音声符号化処理の中で最も処理負荷の重い心理聴覚分析処理を行わない方式

が 3 点提案されている。

【 0 0 1 5 】

1 つ目は各サブバンドの音圧に関わらず、人間の聴覚で聞き取りやすいサブバンドには無条件でビットを割当てする方法であり、場合によっては音圧がほとんどなくてもビットが割当てられる場合が生じる方式である。

【 0 0 1 6 】

2 つ目は人間の聴覚で聞き取りやすいサブバンドかどうかの重み付けと、各サブバンドの音圧から、各サブバンドに割り当てられるビットの比率を求め、この比率に合うようにビットを割り振る手法である。

【 0 0 1 7 】

3 つ目は人間の聴覚で聞き取りやすいサブバンドかどうかの重み付けと、各サブバンドのスケールファクタ値から、ビット割当て情報係数と呼ばれる各サブバンドに対するビット割当て優先順位を求め、優先順位の高いサブバンドから順にビットを割り当てていく手法である。

【 0 0 1 8 】

また、2 5 5 8 9 9 7 号特許公報では各サブバンド信号に対して、2 種類の重み付けをすることで音声符号化処理の負荷を軽減する方式が提案されている。1 つ目の重み付けはサブバンド信号のレベルの対数値に対する重み付けであり、2 つ目は各サブバンド毎に予め定められる重み付けである。1 つ目の重み付けが心理聴覚分析処理に代わるものという位置付けである。

【 0 0 1 9 】

さらに、特開平 1 1 - 3 3 0 9 7 7 号公報では各サブバンドを量子化誤差でランク付けし、量子化誤差が大きくなるサブバンドは符号化せず、量子化誤差の小さいサブバンドにだけビットを与えて符号化する方式が提案されており、音質を保った状態で符号化効率を向上させている。ここではこの方式を、符号化する周波数範囲を適応的に変化させることから「適応スケーラブルコーディング」と呼んでいる。

【 0 0 2 0 】

これら公報記載の技術は、いずれも音声符号化処理の負荷を軽減させるための

ものであるが、低演算量で心理聴覚分析処理を実現することにより、音声符号化処理の負荷を軽減したものではない。

【 0 0 2 1 】

そこで、本発明の目的は、音声符号化処理において低演算量で心理聴覚分析処理を実現することができ、処理負荷を軽減した効率の良い音声符号化環境を実現することができる音声符号化装置及びそれに用いる心理聴覚分析方法を提供することにある。

【 0 0 2 2 】

【課題を解決するための手段】

本発明による音声符号化装置は、入力信号を複数の周波数帯域に分割する分割手段を持ち、前記分割手段で分割された各サブバンド信号を圧縮符号化する音声符号化装置であって、前記各サブバンド信号の各周波数について聴感上の音の大きさが等しい音圧レベルの値を結んだ等ラウドネス曲線に準拠した重み付けを行いかつその重み付けされた量子化誤差が前記各サブバンド信号で均等になるようにビット割当てを行う手段を備えている。

【 0 0 2 3 】

本発明による心理聴覚分析方法は、入力信号を複数の周波数帯域に分割する分割手段を持ち、前記分割手段で分割された各サブバンド信号を圧縮符号化する音声符号化装置の心理聴覚分析方法であって、前記各サブバンド信号の各周波数について聴感上の音の大きさが等しい音圧レベルの値を結んだ等ラウドネス曲線に準拠した重み付けを行いかつその重み付けされた量子化誤差が前記各サブバンド信号で均等になるようにビット割当てを行うステップを備えている。

【 0 0 2 4 】

すなわち、本発明の心理聴覚分析方法は、MPEG (Moving Picture Experts Group) 規格のような人間の聴覚を考慮した音声符号化方法において処理負荷を軽減した効率の良い心理聴覚分析を提供する方法である。

【 0 0 2 5 】

例えば、MPEG規格における心理聴覚分析は人間の聴覚能力の限界やマスキ

ング効果を考慮した上で、各帯域にビットを割当て際の優先順位を決定する手段であり、規格書では聴覚心理モデルと呼び、その処理手順が示されている。人間の聴覚によって聞き取りやすい音声帯域により多くの符号化ビットを割当てることで、再生音質の優れた符号化音声データを取得することができる。

【 0 0 2 6 】

しかしながら、規格書に示された聴覚心理モデルは F F T（高速フーリエ変換）に始まり、F F T で求められた信号に対して最小可聴限界との比較や、さらにはマスキング効果の分析等の処理負荷が重くかつ複雑な演算をする必要がある。

【 0 0 2 7 】

特に、パーソナルコンピュータ等の C P U（中央処理装置）上で動作するソフトウェアによって M P E G 規格による音声符号化装置を実現した場合、聴覚心理モデルの負荷が非常に重く、符号化処理を実現するパーソナルコンピュータ等の性能によっても符号化性能が大きく左右される。その場合、性能の悪い装置で実時間符号化（リアルタイム符号化）処理を施すと、符号化処理が間に合わずに、再生時に音声途切れが生じてしまうこともある。本発明の心理聴覚分析方法は、これらの問題を解決するようにしたことを特徴とする。

【 0 0 2 8 】

より具体的に、本発明の心理聴覚分析方法では、等ラウドネス曲線に準じて各サブバンドの重み付け係数を設定し、加えて各サブバンドの許容量量子化誤差値の初期値を設定する。次に、各サブバンドのスケールファクタ値と、重み付け係数及び許容量量子化誤差値からビット割当て可能な全てのサブバンドに対して量子化ステップ数を算出する。

【 0 0 2 9 】

その後、割当てられた総符号量を算出し、総符号量が許容符号量を超えていた場合に、新たに許容量量子化誤差値を設定し、再び各サブバンドに対して量子化ステップ数を算出する。総符号量が許容符号量以下であった場合には、新たな許容量量子化誤差値を設定した後、その許容量量子化誤差値がビット割当ての収束条件を満たしたかどうかを判断し、満たされていないと判断すると、再び各サブバンドに対して量子化ステップ数を算出する。収束条件を満たしたと判断された場合

には、聴覚分析ビット割当て処理を終了する。

【0030】

従来、聴覚心理モデルでの演算結果を基にビット割当て処理を施しているが、本発明による手法によって各サブバンドの量子化誤差が均等になるようにビット割当てを行うため、聴覚心理モデルを使用せずに符号化することが可能である。

【0031】

また、各サブバンドの重み付け係数を設定する際に、設定された符号化ビットレートを確認し、基準以下のビットレートであると判断した時に、等ラウドネス曲線に準じた各サブバンドの重み付け係数を、このビットレートに応じてさらに重み付けする。これによって、符号化ビットレートに応じた音質を維持し、符号量不足による符号化ノイズの発生も防いだ状態で、幅広い符号化ビットレートに対応することが可能となる。

【0032】

【発明の実施の形態】

次に、本発明の一実施例について図面を参照して説明する。図1は本発明の一実施例による音声符号化装置の構成を示すブロック図である。図1において、音声符号化装置1はサブバンド分析部11と、スケーリング部12と、聴覚分析ビット割当て部13と、量子化部14と、ビットストリーム生成部15とから構成されている。

【0033】

サブバンド分析部11は入力信号を複数の周波数帯域に分割する。スケーリング部12は各サブバンド信号に対して基準値からの倍率であるスケールファクタを計算し、ダイナミックレンジを揃える。

【0034】

聴覚分析ビット割当て部13には本発明の一実施例による心理聴覚分析方法が組込まれている。量子化部14は量子化計算を行う。ビットストリーム生成部15はヘッダや補助情報と共にビット列を形成する。

【0035】

聴覚分析ビット割当て部13は各サブバンド信号に対して等ラウドネス曲線に

準じた重み付けをした後、重み付けされた量子化誤差が各サブバンドで均等になるようにビット割当て量を算出する。

【0036】

また、聴覚分析ビット割当て部13では各サブバンド信号に対して等ラウドネス曲線に準じた重み付けをする他に、符号化ビットレートに応じた重み付けを付加し、重み付けされた量子化誤差が各サブバンドで均等になるようにビット割当て量を算出することもできる。

【0037】

人間には個人差があるものの、実際には同じ音圧レベルを持った信号であっても、その周波数によって聴感上の音の大きさが異なる。純音の各周波数について、聴感上の音の大きさが等しい音圧レベルの値を結んだ曲線を等ラウドネス曲線、または音の大きさの等感曲線と呼ぶ。つまり、周波数に関わらず全て同一の音圧レベルを持った音声信号であったとしても、聴感上は異なる音の大きさに聞こえるということである。

【0038】

この曲線から、人間が最も知覚し易い周波数は4kHz付近であり、この4kHzを中心にして高周波数／低周波数になるにしたがい、知覚しにくくなる。等ラウドネス曲線については「音響振動工学」（西山他、コロナ社、昭和54年4月、P23）等に詳しく述べられている。

【0039】

図2は図1の聴覚分析ビット割当て部13の動作を示すフローチャートであり、図3は本発明の一実施例における等ラウドネス曲線に準拠したサブバンド単位の重み付けテーブルの一例を示す図であり、図4はMPEG1/Audioレイヤ1符号化方式における量子化ステップ数と割当てビット数との関係を示す図である。これら図1～図4を参照して本発明の一実施例による心理聴覚分析方法について説明する。尚、本発明の一実施例ではMPEG1/Audioレイヤ1を例として説明する。

【0040】

16ビット直線量子化された入力信号はサブバンド分析部11で32帯域のサ

ブバンド信号に分割される。各サブバンド当たり12サンプルで、合計384サンプル単位で以降の処理が実行される。この32帯域に分割された各サブバンド信号のダイナミックレンジを揃えるため、スケーリング部12では最大振幅が1.0になるように正規化し、その倍率であるスケールファクタを各サブバンド単位で算出する。

【0041】

次に、聴覚分析ビット割り当て部13で各サブバンドに対するビット割当て量を決定する。最初に、初期設定を行う（図2ステップS1）。この初期設定ではまず予め各サブバンドに対する重み付け係数を決定しておく。この重み付け係数は等ラウドネス曲線に準拠して決定される。つまり、人間の最も知覚しやすい周波数帯域を持つサブバンドに、最も多くビットが割当てられるよう重み付け係数を決定することとなる。

【0042】

等ラウドネス曲線によれば、4kHz付近が最も知覚しやすい帯域であることを判断することができる。今回は、係数値が大きくなるほど当該サブバンドへのビット割当て優先度が低くなるものとし、最もビット割当て優先度が高い場合の係数値を1.0としている。

【0043】

ここで、基本概念について説明する。各サブバンドにおけるスケールファクタを $scale(s_b)$ 、量子化ステップ数を $qsteps(s_b)$ とすると、量子化誤差 $qerr(s_b)$ は、

$$qerr(s_b) = scale(s_b) / qsteps(s_b) \\ (s_b = 0, 1, 2, \dots, 31)$$

となる。

【0044】

また、各サブバンドに対する重み付け係数を $weight(s_b)$ とした場合、重み付け量子化誤差 $wqerr(s_b)$ は、

$$wqerr(s_b) = qerr(s_b) \times weight(s_b) \\ (s_b = 0, 1, 2, \dots, 31)$$

で表される。

【0045】

この重み付け量子化誤差 $w_{qerr}(sb)$ が各サブバンドで等しくなり、かつ $w_{qerr}(sb)$ が許容符号量内で最小値になるように $qsteps(sb)$ を制御することによって、人間の聴覚心理を利用したビット割当てを行うことになる。

【0046】

次に、許容量子化誤差の初期値を設定する。許容量子化誤差とは各サブバンドにおけるスケールファクタの内の最大値を、各サブバンドに割当て可能な仮の最大量子化ステップ数で除算したものであり、この時点で最小の量子化誤差値ということになる。

【0047】

スケールファクタの最大値を max_scale とし、割当て可能な仮の最大量子化ステップ数を「255」とした時、許容量子化誤差 err_thr の初期値は、

$$err_thr = max_scale / 255$$

で与えられる。

【0048】

量子化ステップ数とは何段階で量子化するかを示すものであり、MPEG1/Audioレイヤ1では全て2のべき乗より1小さい値で示され、最大値は「32767」で、最小値は「3」である。また、量子化しない場合には量子化ステップ数に「0」が与えられる。

【0049】

さらに、MPEG1/Audioレイヤ1の場合、各サブバンドに対して実際に割当て可能な最大量子化ステップ数は「32767」と規定されており、この場合に最も誤差が少ない量子化が可能ということになる。

【0050】

一方、最小量子化ステップ値「3」の場合には、最も誤差が大きい量子化ということになる。このことから、初期段階での最も細かい量子化誤差 err_th

r_min と、最も粗い量子化誤差 err_thr_max とは、

$$err_thr_min = max_scale / 32767$$

$$err_thr_max = max_scale / 3$$

という式のように示される。これらの式は総符号量算出の際に、量子化誤差が規定内に収まったかどうかの判断に使用される。

【0051】

以上で初期設定が終了し、次に各サブバンドの量子化ステップ数が算出される(図2ステップS2)。各サブバンドの量子化ステップ数 $q_steps(s b)$ は、

$$q_steps(s b) = scale(s b) \times weight(s b) / err_thr$$

$$(s b = 0, 1, \dots, 31)$$

という式で求められる。

【0052】

ここで、求められた量子化ステップ数 $q_steps(s b)$ を、MPEG1/Audioレイヤ1で規定されている量子化ステップ数に丸め込む必要がある。図4に規定されている量子化ビット数と対応する量子化ステップ数との関係を示す。本例では最寄りの量子化ステップ数に切り下げることとしている。

【0053】

次に、各サブバンドに割当てられた量子化ステップ数から、対応する量子化ビット数を図4にしたがって取得し、さらにサイド情報やヘッダ情報等のMPEG1/Audioビットストリーム構成に必要なビット数を加算した上で、総符号量を取得する(図2ステップS3)。

【0054】

この総符号量を符号化ビットレートによって決定される実際に割当て可能な許容符号量と比較する。ここで、総符号量が許容符号量を超えている場合(図2ステップS4)、現在の許容量子化誤差 err_thr が細かすぎたものと判断することができるため、許容量子化誤差 err_thr を粗くする方向で更新する(図2ステップS5)。

【0055】

許容量子化誤差 err_thr の更新は次のように実行する。まず、現在の許容量子化誤差 err_thr は、新たな最も細かい量子化誤差 err_thr_min として保存する。つまり、

$$err_thr_min = err_thr$$

となる。

【0056】

この後、新たな許容量子化誤差値を、

$$err_thr = (err_thr + err_thr_max) / 2$$

という式で算出する。このようにして許容量子化誤差を更新した後、再度各サブバンドの量子化ステップ数を算出する（図2ステップS2）。

【0057】

一方、総符号量が許容符号量以下であると判断された場合（図2ステップS4）、現在の許容量子化誤差が粗すぎたものと判断することができるため、許容量子化誤差を細かくする方向で更新する（図2ステップS6）。

【0058】

許容量子化誤差 err_thr の更新は次のように実行する。まず、現在の許容量子化誤差 err_thr を、新たな最も粗い量子化誤差 err_thr_max として保存する。つまり、

$$err_thr_max = err_thr$$

となる。

【0059】

この後、新たな許容量子化誤差値を、

$$err_thr = (err_thr + err_thr_min) / 2$$

という式で算出する。

【0060】

ここで、新たな許容量子化誤差値を基にビット割当て処理が収束したかどうかの判断をする。この場合、

$$err_thr / err_thr_max > 0.9$$

という式の条件が満たされた時に、ビット割当て処理が収束したとみなし、処理を終了する（図2ステップS7）。

【0061】

一方、上記の式の条件が満たされなかった時には、まだビット割当て処理が収束していないとみなし、この更新した許容量子化誤差 err_thr を使用して、再度各サブバンドの量子化ステップ数を算出する（図2ステップS2）。

【0062】

次に、量子化部4で対称零表現による線形量子化器を用いて各サブバンド信号を量子化した後、ビットストリーム生成部5でヘッダ情報及びサイド情報と共にビット列を形成し、符号化処理を終了する。

【0063】

上記のように、本実施例によるビット割当て手法によって、規格書に示された心理聴覚モデルを使用したビット割当て手法のように、FFT（高速フーリエ変換）やマスキング効果の分析等の処理負荷の重い複雑な計算をすることなく、ビット割当て処理を行うことができるため、符号化処理負荷を軽減することができる。

【0064】

図5は本発明の他の実施例における重み付けテーブルを符号化ビットレートに対応した重み付けテーブルに更新する手法を示すフローチャートであり、図6は本発明の他の実施例における符号化ビットレートに対応したサブバンド単位の重み付けテーブルの一例を示す図であり、図7は本発明の他の実施例における推奨ビットレート未満の場合の聴覚分析ビット割当て部13の動作を示すフローチャートである。

【0065】

本発明の他の実施例による音声符号化装置は聴覚分析ビット割当て部13の動作が異なる以外は図1に示す本発明の一実施例による音声符号化装置1と同様の構成となっているので、その説明は省略する。以下、これら図1及び図5～図7を参照して本発明の他の実施例について説明する。

【0066】

本発明の一実施例では全てのサブバンドに対してビットを割当てする前提で等ラウドネス曲線に準拠した重み付けテーブルを作り、ビット割当てを行っているが、符号化ビットレートが小さい場合には、特にターゲットビットレートと呼ばれる推奨ビットレート未満の場合には、符号化ビットレートが大きい場合と同様の重み付けでは割当てビット数が不足し、音質の劣化や符号化ノイズ発生の原因となることがある。

【 0 0 6 7 】

このような場合、高音域側のサブバンドに対するビット割当て優先度を下げ、人間が知覚しやすい周波数帯に対してより多くのビットが割当てられるようにすることで、各符号化ビットレートに見合った音質を維持するとともに、符号化ノイズの発生を抑えることができる。以下、符号化ビットレートがターゲットビットレート未満であった場合について説明する。

【 0 0 6 8 】

まず、各サブバンドへの重み付け係数を算出する（図 7 ステップ S 2 1）。この各サブバンドへの重み付け係数の算出では最初に、使用者から設定された符号化ビットレートを確認し（図 5 ステップ S 1 1）、その符号化ビットレートがターゲットビットレート未満であるかどうかの判断を行う。ターゲットビットレート以上であると判断された場合には（図 5 ステップ S 1 2）、図 3 に示す等ラウドネス曲線に準拠した重み付けテーブルをそのまま使用する。

【 0 0 6 9 】

一方、符号化ビットレートがターゲットビットレート未満であると判断された場合には（図 5 ステップ S 1 2）、図 6 に示すビットレート対応重み付け係数と図 3 に示す等ラウドネス曲線に準拠した重み付け係数とを使用し、新たな重み付け係数を算出する（図 5 ステップ S 1 3）。

【 0 0 7 0 】

等ラウドネス曲線に準拠した重み付け係数を $w e i g h t (s b)$ 、ビットレート対応重み付け係数を $w e i g h t _ b r (s b)$ とすると、新たな重み付け係数 $w e i g h t _ n e w (s b)$ は、

$$w e i g h t _ n e w (s b)$$

$$= \text{weight}(sb) \times \text{weight_br}(sb) \\ (sb = 0, 1, 2, \dots, 31)$$

という式で求められる。

【0071】

次に、ビット割当て処理を行うにあたっての初期設定を行う（図7ステップS22）。符号化ビットレートがターゲットビットレート以上ならば、重み付け係数には $\text{weight}(sb)$ を使用し、ターゲットビットレート未満であれば、 $\text{weight_new}(sb)$ を用いる。

【0072】

初期設定手法については本発明の一実施例でのステップS1と同様に処理される。また、以降のビット割当て処理本体（図7ステップS23～S28の処理）についても、本発明の一実施例の処理（図2のステップS2～S7の処理）と同様に処理され、ビット割当て処理が終了される。

【0073】

上記のように、各サブバンドに対して符号化ビットレートに応じた重み付けも加えることによって、符号化ビットレートに見合った音質を維持するとともに、符号化ノイズ発生を抑えた音声符号化を行うことができる。

【0074】

このように、従来の心理聴覚モデルを使用したビット割当て処理を行うことなく、各サブバンド信号に対して等ラウドネス曲線に準拠した重み付けを行うとともに、重み付けされた量子化誤差が各サブバンドで均等になるようにビット割当てを算出することによって、心理聴覚処理を伴った音声符号化処理において、符号化品質を維持した状態で符号化処理負荷を軽減することができる。

【0075】

また、各サブバンドに対して等ラウドネス曲線に準拠した重み付け係数テーブルを持たせる他に、符号化ビットレートに対応した重み付けテーブルを持ち、双方を参照することで符号化ビットレートに応じたビット割当てを行うことによって、心理聴覚処理を伴った音声符号化処理において、符号化ビットレートを低くする方向に変更しても、その符号化ビットレートに応じた音質を維持し、符号量

不足による符号化ノイズ発生をも抑えた音声符号化を行うことができる。

【0076】

尚、本発明の一実施例及び他の実施例ではMPEG1/Audioレイヤ1の場合について述べたが、聴覚心理モデルを用いたビット割当て手段を持つ他の音声符号化方式に対しても本発明を適用することが可能である。この音声符号化方式としては、例えばMPEG1/Audioレイヤ2、MPEG1/Audioレイヤ3、MPEG2/Audio AAC等がある。

【0077】

また、本発明の他の実施例で説明した符号化ビットレートに対応した重み付けテーブルを符号化ビットレートに応じて複数個用意し、適宜使用するテーブルを換えることで、より音質を重視した音声符号化を行うことも可能である。

【0078】

【発明の効果】

以上説明したように本発明によれば、入力信号を複数の周波数帯域に分割する分割手段を持ち、分割手段で分割された各サブバンド信号を圧縮符号化する音声符号化装置において、各サブバンド信号の各周波数について聴感上の音の大きさが等しい音圧レベルの値を結んだ等ラウドネス曲線に準拠した重み付けを行いかつその重み付けされた量子化誤差が各サブバンド信号で均等になるようにビット割当てを行うことによって、音声符号化処理において低演算量で心理聴覚分析処理を実現することができ、処理負荷を軽減した効率の良い音声符号化環境を実現することができるという効果がある。

【0079】

また、本発明によれば、各サブバンド信号を等ラウドネス曲線に準拠した重み付けを行うことに加え、符号化ビットレートに対応した重み付けも行うことで、符号化ビットレートを低くする方向に変更しても、符号化ビットレートに応じた音質を維持するとともに、符号量不足によるノイズ発生を抑えた音声符号化環境をも実現することができるという効果がある。

【図面の簡単な説明】

【図1】

本発明の一実施例による音声符号化装置の構成を示すブロック図である。

【図 2】

図 1 の聴覚分析ビット割当て部の動作を示すフローチャートである。

【図 3】

本発明の一実施例における等ラウドネス曲線に準拠したサブバンド単位の重み付けテーブルの一例を示す図である。

【図 4】

MPEG1/Audioレイヤ1符号化方式における量子化ステップ数と割当てビット数との関係を示す図である。

【図 5】

本発明の一実施例における重み付けテーブルを符号化ビットレートに対応した重み付けテーブルに更新する手法を示すフローチャートである。

【図 6】

本発明の一実施例における符号化ビットレートに対応したサブバンド単位の重み付けテーブルの一例を示す図である。

【図 7】

本発明の一実施例における推奨ビットレート未満の場合の聴覚分析ビット割当て部の動作を示すフローチャートである。

【図 8】

MPEG1/Audioレイヤ1符号化装置の構成を示すブロック図である。

【図 9】

図 8 のビット割当て部の動作を示すフローチャートである。

【図 10】

図 8 の心理聴覚分析部の構成を示すブロック図である。

【符号の説明】

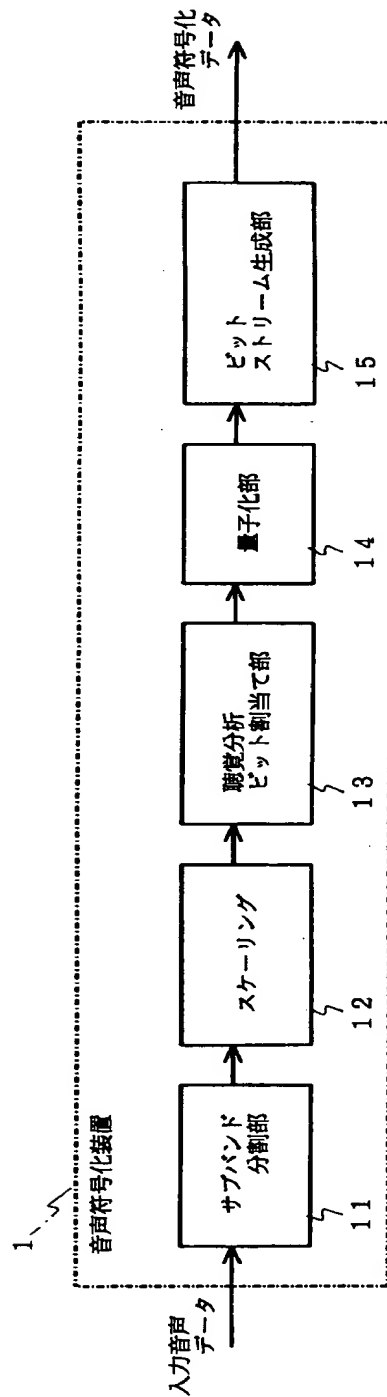
- 1 音声符号化装置
- 1 1 サブバンド分析部
- 1 2 スケーリング部
- 1 3 聴覚分析ビット割当て部

1 4 量子化部

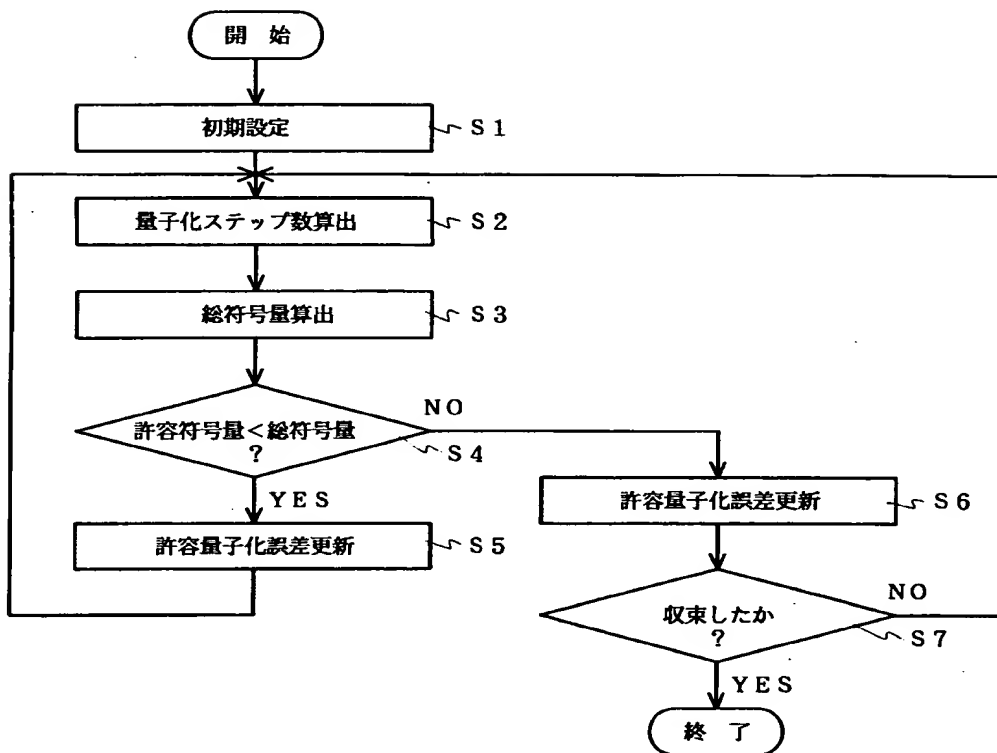
1 5 ビットストリーム生成部

【書類名】 図面

【図 1】



【図 2】



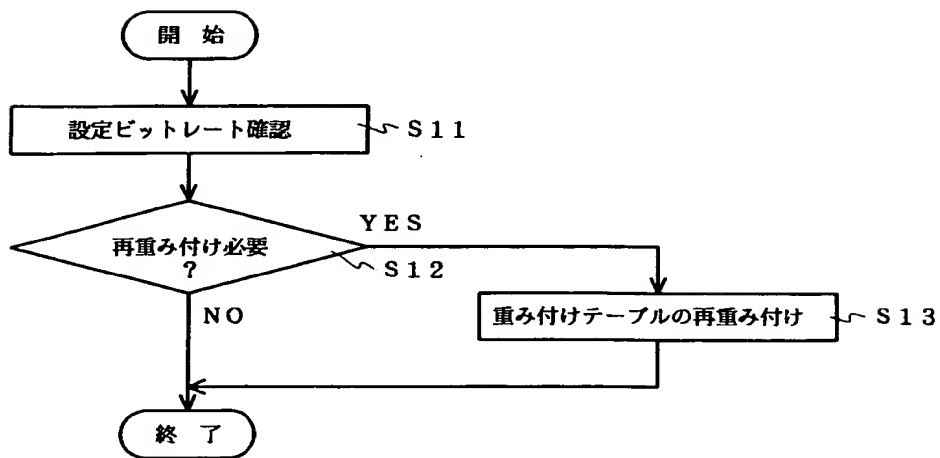
【図 3】

サブバンド (s b)	重み付け係数 w e i g h t
0	1. 3
1	1. 2
2	1. 1
3	1. 1
4	1. 1
5	1. 0
6	1. 0
7	1. 2
8	1. 2
9	1. 2
10	1. 2
11	1. 2
12	1. 3
13	1. 3
14	1. 5
15	1. 5
16	1. 5
17	1. 5
18	1. 7
19	1. 9
20	4. 5
21	4. 5
22	4. 5
23	4. 5
24	4. 5
25	5. 0
26	5. 0
27	5. 0
28	5. 0
29	5. 0
30	5. 0
31	5. 0

【図 4】

量子化ステップ	ビット数
0	0
3	2
7	3
15	4
31	5
63	6
127	7
255	8
511	9
1023	10
2047	11
4095	12
8191	13
16383	14
32767	15

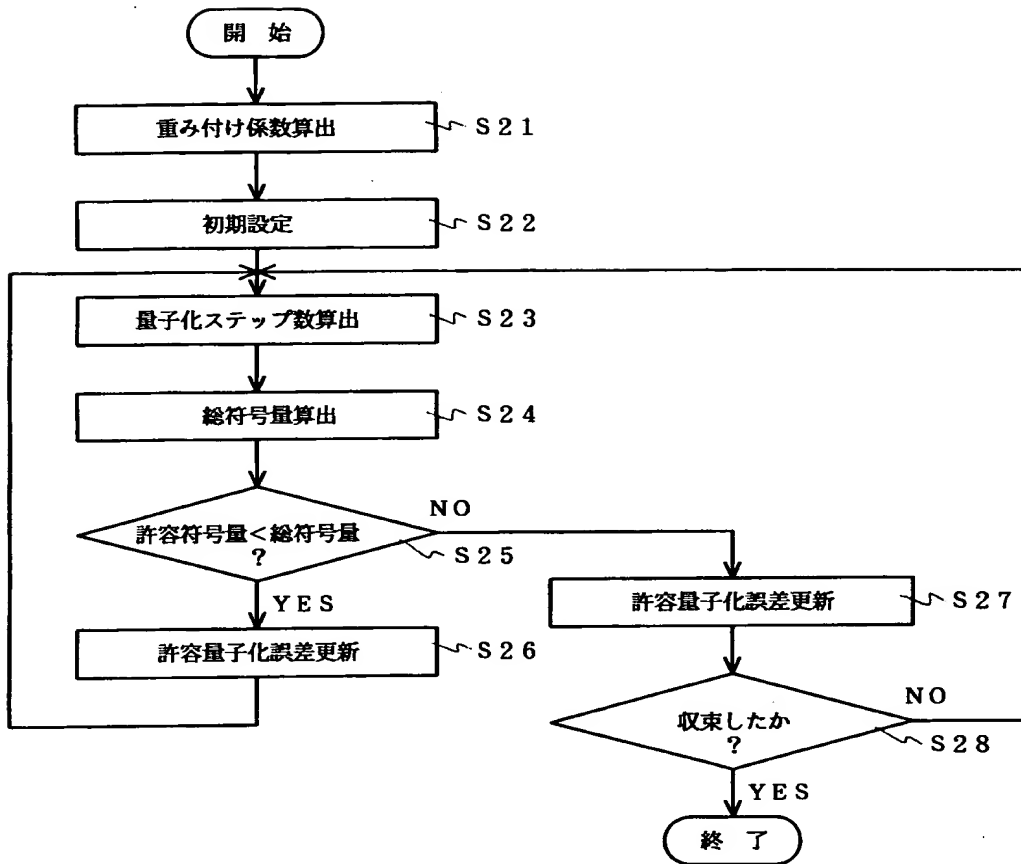
【図 5】



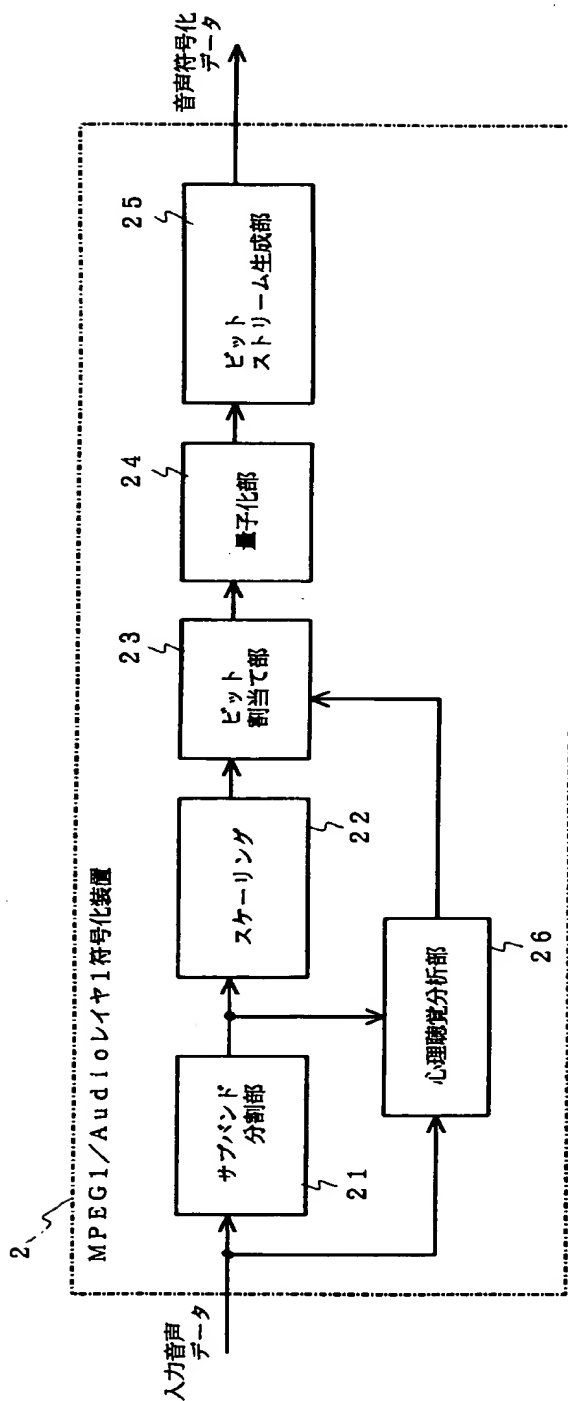
【図 6】

サブバンド (s b)	重み付け係数weight
0	1. 0
1	1. 0
2	1. 0
3	1. 0
4	1. 0
5	1. 0
6	1. 1
7	1. 2
8	1. 5
9	1. 5
10	1. 5
11	1. 5
12	1. 5
13	1. 5
14	1. 5
15	1. 5
16	2. 5
17	2. 5
18	2. 5
19	2. 5
20	2. 5
21	2. 5
22	2. 5
23	2. 5
24	3. 0
25	3. 0
26	3. 0
27	3. 0
28	3. 0
29	3. 0
30	3. 0
31	3. 0

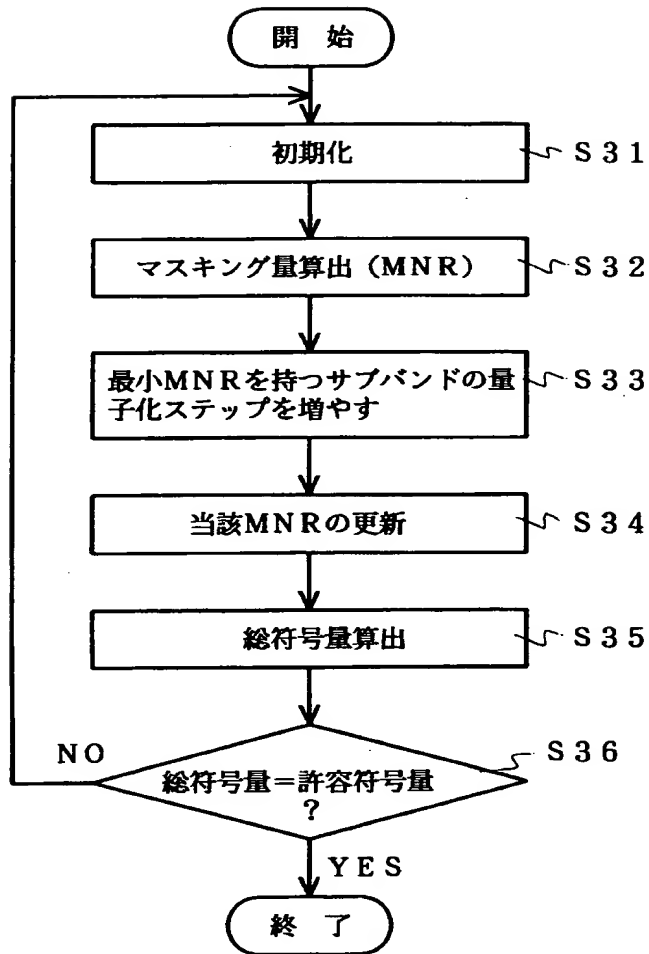
【図 7】



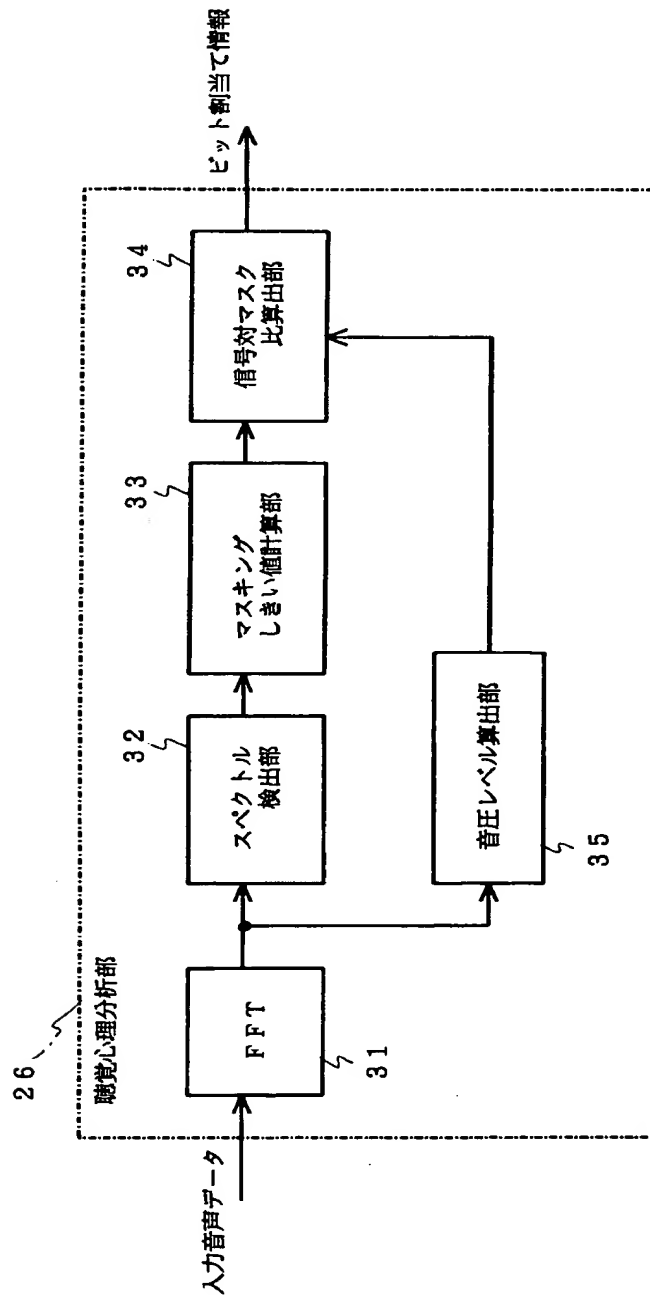
【図 8】



【図9】



【図10】



【書類名】 要約書

【要約】

【課題】 音声符号化処理において低演算量で心理聴覚分析処理を実現し、処理負荷を軽減した効率の良い音声符号化環境を実現可能な音声符号化装置を提供する。

【解決手段】 サブバンド分析部 1 1 は入力信号を複数の周波数帯域に分割する。スケーリング部 1 2 は各サブバンド信号に対して基準値からの倍率であるスケールファクタを計算し、ダイナミックレンジを揃える。聴覚分析ビット割当て部 1 3 は各サブバンド信号に対して等ラウドネス曲線に準じた重み付けをした後、重み付けされた量子化誤差が各サブバンドで均等になるようにビット割当て量を算出する。量子化部 1 4 は量子化計算を行う。ビットストリーム生成部 1 5 はヘッダや補助情報と共にビット列を形成する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [000004237]

1. 変更年月日	1990年 8月29日
[変更理由]	新規登録
住 所	東京都港区芝五丁目7番1号
氏 名	日本電気株式会社